Astronomical Image Colorization and Super Resolution using Residual Encoder Networks and GANs

Gao Xian Peh, Kai Marshland
gxpeh@stanford.edu, kaimarsh@stanford.edu
Computer Science Department, Stanford University

Prediction

A huge number of raw images lie unprocessed and unseen in the Hubble Legacy Archives. These raw images are typically low-resolution, black and white, and unfit to be shared with the world. It take tens, or even hundreds of hours to process each one so that astronomers can more easily distinguish objects and so that the public can get breathtaking view of the universe. Random and systematic noise from the sensors in the telescope, changing optical characteristics in the system, and noise from other bodies in the universe all make processing necessary and difficult. NASA specifically asks the public for help in processing images from the Hubble Legacy Archive, which is why we set out to do this task of artificial colourization and super resolution.

Data & Features

We have two types of data; millions of raw images from the Hubble Legacy Archive and approximately 10,000 high-resolution images generated by scraping and augmenting images from the Hubble Heritage project and the main Hubble website. The raw images are black and white and low resolution. The high-resolution images are in colour and have been painstakingly generated by humans; because there were so few, we had to augment the dataset with techniques like cropping. A great deal of scraping and pre-processing went into creating a dataset suitable for use in our models. We fed the image data directly into our models, without further feature extraction.

Models

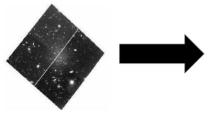
For colourization, we use a residual encoder network that takes in the higher layers of a pre-trained VGG-16 model. It passes these through a series of batch normalizations and convolutions to generate colour values.

For image super-resolution, we use a SRGAN which uses a SRResNet as its generator and an objective function which is the minimax game: $\min_{\theta_G} \max_{\theta_D} \mathbb{E}_{I^{HR} \sim p_{train}(I^{HR})}[\log D_{\theta_D}(I^{HR})] + \mathbb{E}_{I^{LR} \sim p_G(I^{LR})}[\log(1 - D_{\theta_D}(G_{\theta_G}(I^{LR})))]$
We use the perceptual loss function which is a weighted sum of a content loss (VGG loss) and an adversarial loss component.

$$l^{SR}_{VGG_{i,j}} = \frac{1}{W_{i,j}H_{i,j}} \sum_{x=1}^{W_{i,j}} \sum_{y=1}^{H_{i,j}} (\phi_{i,j}(I^{HR})_{x,y} - \phi_{i,j}(G_{\theta_G}(I^{LR}))_{x,y})^2 \qquad l^{SR}_{Gen} = \sum_{n=1}^{N} -\log D_{\theta_D}(G_{\theta_G}(I^{LR}))$$

Results

Raw Hubble Images (left) are transformed into coloured, high-resolution images suitable for use (right)



The super-resolution aspect dramatically outperforms the bi-cubic baseline visually, though the PNSR metric indicates otherwise

Bi-cubic baseline *SRGAN*

PNSR is calculated as:

$$10 \log_{10} 255^2 / \text{MSE}$$

This seems misleading, given that visually the GAN is clearly better.

Model	Mean Squared Error	Peak Signal to Noise Ratio	Inception Score
Image Colourization + SRGAN	201.769	27.393	3.262 (Std Dev: 0.327)
Bicubic Interpolation on Coloured Images	180.271	38.092	-
Original Images	-	-	3.447 (Std Dev: 0.206)

Qualitatively, our Residual Encoder Image Colourization + Super-Resolution SRGAN model performed admirably, especially on the raw, unprocessed Hubble images we care most about.

Discussion

Though not yet perfect, our project demonstrates that automatic colourization and super-resolution produces images far better than the raw data collected by Hubble.

Qualitatively, we observed high-quality results, for colourization and super-resolution individually as well combined. There are undeniably flaws; colours tend toward the red and green, with almost no blues. Colours are also somewhat more washed out. The quantitative errors may actually underestimate the quality of the results, given that *all* colours are artificial, and what matters is much more the qualitative visuals of the result, rather than a true distance. The super-resolution aspect of the project also performed well; visually, it looks significantly sharper than the bi-cubic baseline. In this case, it seems that PSNR is a misleading metric; our SRGAN approach gives dramatically better results than the bi-cubic baseline, despite receiving a lower score.

Future

In the future, our images could be directly studied by astronomers or an image stitching algorithm can be applied on our archive images to generate large-scale astronomical images for scientific study. On the technical front, colourization could be improved with further experimentation with the loss function, like trying to increase loss for low saturations, a common problem we saw. We could also attempt to use a single GAN for both colourization and super-resolution.

References

R. Dahl, "Automatic colorization."
K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014.
I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio,"Generative adversarial nets," inAdvances in neural information processing systems, pp. 2672–2680, 2014.
C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz,Z. Wang,et al., "Photo-realistic single image super-resolution using a generative adversarial network," inProceedings of the IEEE conference on computer vision and pattern recognition, pp. 4681–4690, 2017

Key references only. Full list of references available in paper.